# A Testbed Demonstrating Optical IP Switching (OIS) in Disaggregated Network Architectures

Marco Ruffini, Donal O'Mahony, Linda Doyle.
Centre for Telecommunication Value Chain Research,
University of Dublin,
Trinity College, Dublin 2, Ireland.

**Abstract:**

*The lack of a unified control plane does not allow current optical networks to dynamically provision new optical paths. The IETF standardization body has proposed the Generalized Multi-Protocol Label Switching standard as a solution to this problem. Their efforts however focus on the creation of end-to-end optical trails. This concept is convenient for major service providers, whose network may span a large area, but may not suit smaller network operators. We believe that the future trend of the telecommunication industry is made up of a global network formed by many inter-linked operators of small and medium size. A novel type of network based on a distributed and disaggregated architecture seems to be the best solution to dynamic optical path establishment.*

*In this paper we present an optical testbed implementing the concept of Optical IP Switching. Central to the system is the optical router that creates new lightpaths depending on encountered traffic flows. The decision making process is completely distributed, and perfectly fits the disaggregated network view. The testbed currently links two universities in Dublin, and may in future be extended to join other existing research networks.*

## 1    Introduction:

Traditionally the provisioning of new optical paths is a lengthy process. In fact the time needed to setup an OC-3 or OC-12 optical connection can be up to 6 months in many locations [1].

This problem is very well-known both to the network customers and providers and is experienced by the former as a delay in fulfilling their bandwidth ambitions and by the latter as a loss of new revenue opportunities. It is interesting to analyze where this latency time originates.

Where there is not enough installed capacity the laying of a new optical fiber would require a set up time much longer than 6 months.

Fortunately this is not the usual case, as recent advances in Wavelength Division Multiplexing mean that the fiber already in place can accommodate many times more bandwidth than it previously did.

The difficulty network providers have is to coordinate and organize their own network elements (not to mention the problems in organizing a connection through different operators).

The control plane is the set of entities that should control and organize the links between the network elements, but current standards are still works in progress.
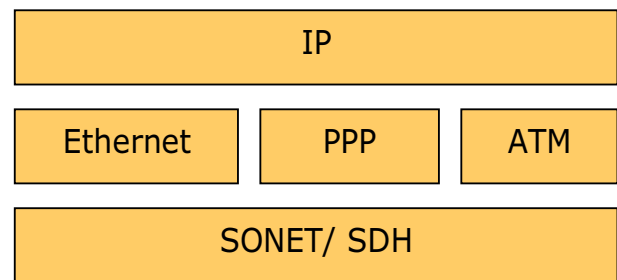


**Figure 1 Protocol stack of an IP over Sonet optical network.**

If we look closely at the typical protocol structure of a packet switched optical architecture (Figure 1), we can see that it is made up of layers that use different and independent technologies: the packet oriented IP network is implemented on top of a circuit switched SONET/SDH optical network, and linked together by an ATM, Ethernet or Point-to-Point protocol layers.

While the IP layer has a routing mechanism that allows automatic neighbour discovery and path computation, the optical layer lacks such capabilities: this is inconsistent with the idea of a fully automated network.

In the absence of a broadly accepted control plane, different vendors have developed their own means for organizing their network elements, which in most cases does not represent either a complete or interoperable solution.

Aware of this condition network service providers and vendors have been working hard in the past few years to fill this gap by pushing three of the most relevant standardization bodies to develop a standard for the optical control plane.

The ITU-T (International Telecommunication Union-Telecommunication Standardization Sector) has worked on the ASON architecture (Automatically Switched Optical Network), which describes the basic functional requirements of the optical network and defines the principles for the UNI (User-Network Interface) and NNI (Network-Network Interface). The IETF (Internet Engineering Task Force) is developing the GMPLS architecture (Generalized Multi-

Protocol Label Switching) which addresses link management, topology discovery, connection provisioning and the protection and restoration issues. Finally the OIF (Optical Internetworking Forum), focuses its attention on the actual implementation of the control interfaces (UNI and NNI) defined in the ASON: their efforts culminated in the OIF World Interoperability Demo, which in 2004 brought different vendors and carriers together for a practical demonstration of the interoperability between their equipments [3].

These standardization bodies are collaborating to produce a set of compatible and interoperable standards [2], [4].

Once completed and fully implemented these standards should accelerate and facilitate the provisioning of optical paths, by automating tasks like topology discovery and connection provisioning that currently require large amount of human intervention.

The GMPLS standardization however seems to satisfy only the needs of the major service providers, whose network may span a large area.

If we have a closer look at GMPLS, we realize that it mainly focuses on the capability of reserving end-to-end optical paths within a unique domain by using one of the two competing signalling protocols RSVP-TE [5] or CR-LDP [6]. In both scenarios a user forwards the request for a dedicated bandwidth to a network operator using a User-Network interface. The GMPLS control plane of the operator then configures an optical path with the required characteristics. Should the connection span different providers, things would be complicated by the need for inter-domain routing (which is not yet well covered by the standardization activity). Moreover the interaction between users and providers involves management and administrative overheads that may still constitute the main element of the delay.

## 2    Towards disaggregation:

Even today it is unlikely that a path can be found from one part of the world to another that only crosses a single network domain.

Moreover we have observed that the demonopolization of the telecommunication industry has favoured, in many countries, the growth of a more competitive market. This tendency was confirmed by recent reports showing that the market share of the top five telecom companies decreased from 70% to 53% in Europe in the past two years, in favour of smaller and new operators [7].

Figure 2 shows the architecture of a disaggregated network, where each operator may manage a small part of the complete picture. In such a multi-provider domain, we believe that it is essential for each router to have the capability to autonomously trigger the creation of new optical light paths.

In contrast to GMPLS what we propose is not a protocol that allows users to reserve bandwidth on a dedicated path, but an optical router with new functionalities. Such a device will switch all the packets belonging to the same flow at the optical

level, saving the processing power needed to route each packet at the network level.
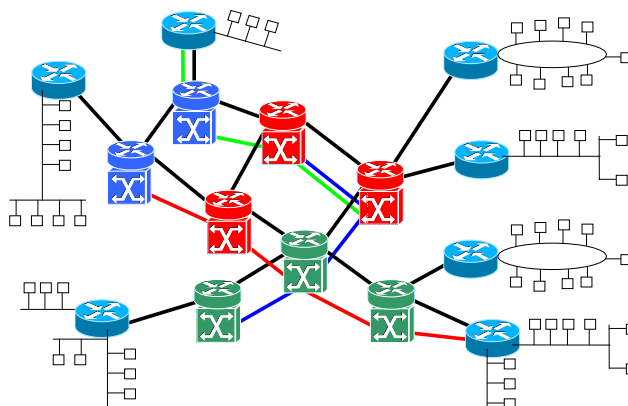


**Figure 2 Disaggregated network view of an IP-switched optical network.**

The idea of having flows bypassing the IP layer is not new and was first introduced in the electrical domain as a method to combine ATM and IP technologies [8],[9]. Subsequently, with the introduction of photonic switches the idea was extended into the optical domain [10],[11], but then the efforts have been directed towards a user-oriented "on-demand" optical provisioning, leading to the development of the ASON and GMPLS standards. The Hikari router [12],[13] can be considered an implementation of that idea in the optical domain. This router, developed by NTT in Japan, allocates new optical paths in response to requests for new MPLS trails. When the existing optical channels do not have enough bandwidth to carry new switched paths a new optical trail is originated on a spare wavelength to accommodate the new traffic.

In this case the establishment of new optical trails depends on incoming MPLS path requests. However no standard currently defines which condition should trigger the creation of an MPLS trail: the OSPF-xTE (the traffic engineering extension of the OSPF protocol) is still a work-in-progress. As a result there is at present no automatic creation of those label switched paths. Currently the human intervention is the intelligence that creates the switched paths: for example a network administrator could trigger the creation of a trail, either following a precise user request or for optimization purposes like traffic balancing.

### 2.1    Optical IP Switching (OIS):

In this paper we propose a novel optical architecture which we call "Optical IP Switching" (OIS), based on an automatic, flow-based method of lightpath creation. The optical router monitors IP traffic and if a flow appears with specific characteristics the router establishes an optical cut-through path between its upstream and downstream neighbours,

requesting the upstream node to place all the packets belonging to the flow into the new path. The newly generated trail bypasses the IP layer of the router, as the packets transparently flow from the upstream to the downstream neighbour.

We want to focus the readers' attention on the fact that the path created does not accommodate a flow from its source to its destination. The initial cut-through path in fact only involves three adjacent nodes. Following a similar procedure the path can then be extended to more than three nodes, but this decision is always autonomously taken by each router and depends on the traffic encountered and on the resources locally available.

This represents a unique feature of our approach and is totally in line with the disaggregated network concept.

The benefits of using IP switching either at an electrical or optical level have already been demonstrated: the bypass of the IP layer saves processing and memory resources of the router, enabling much greater network throughput and less latency.

From a resources conservation point of view the optical cut-through path gets around the problem of limited processing power at the routing layer by exploiting unused wavelengths present on the many dark fibers already in place.

The novelty of our idea is that we want the optical paths to be dynamically and automatically created, depending on traffic flows encountered at the routers.

Whether this choice is economically advantageous or not mainly depends on how well we manage to exploit the dynamically created optical paths.

This depends on the characteristics of the traffic traversing the router: the ideal situation is when a small number of flows carry most of the network traffic (those flows are sometimes dubbed "Elephants"). The small number of flows would be consistent with the fact that today's optical switches have a limited number of ports (typically 256 or less), and carrying a large amount of traffic would best exploit the very coarse granularity of an optical light path. This assumption seems to be confirmed by recent analysis showing that Internet traffic is characterized by a "heavy tailed" distribution, where very few flows carry a high percentage of data traffic [14],[15].

Moreover when there are not many "Elephants", smaller flows can be aggregated together to make use of the available bandwidth. All the flows that share a common path could be redirected over the same dedicated lighpath.

## 2.2 Architectures Comparison

The main focus of the previously mentioned optical standards is to offer bandwidth-on-demand services to customers. In order to generally improve routing performance however, protocols like GMPLS need the support of a traffic engineering service with global view of the network traffic.

The architecture we propose instead achieves the same goal by continuously and automatically re-shaping optical paths within the network to optimally support the offered packet load.

A comparison can be also made between OIS and Optical Burst Switching (OBS) [16], [17] as both the architectures aim at speeding up the routing process with traffic-driven lightpath creation mechanisms.

Optical Burst switching is a technique that raised a lot of interest all over the world in the past few years and consists on aggregating packets with similar destination at an edge node; after collecting a sufficient amount of packets, the node triggers the creation of an optical point-to-point connection where the previously aggregated packets are sent as a burst to the network egress point. The network resources are released straight after the burst has been transferred.

The differences between OBS and OIS all stem from the fact that the former handles data in terms of packets while the latter in terms of flows.

Optical IP switching is best suited to follow the slower dynamics of larger flows in a network, where fine granularity is not an issue and data can be handled at a flow level. On the other hand Burst switching is better at handling more dynamic traffic patterns, as it can follow data with granularities of the order of groups of packets.

Due to the heavy tailed distribution of the Internet traffic however we find that a flow-based approach better fits the underlying traffic and could have higher impact on a larger fraction of the Internet traffic, when compared to an OBS-like approach that treats the traffic like completely independent bursts.

OIS also presents other advantages in term of network latency and failure resistance. While OBS inherently introduces latency when accumulating the packets that will constitute the burst, the jitter introduced by the transition from routed to switched flows in OIS tends to be minimized as the router targets larger "Elephant" flows.

In terms of network failures, an OBS network has a certain probability of loosing an entire burst if the required resources are not available along a path at a given instant. By contrast OIS can take advantage of an existing IP network to route the packets when no enough resources are available to switch a new flow on a dedicated path.

Finally OBS requires much faster switching times (in the order of microseconds or less) that current devices are not able to provide.

## 2.3 Network Overview:

Figure 2 gives an overview of the architecture of an optical IP-switched network.

The routers with photonic switching capabilities are those in the central part of the figure, representing the backbone of the network and are linked to each other by WDM connections (groups of nodes may belong to different service providers).

The routers without switching capabilities are at the edge of the network and can utilize one or more wavelengths to connect to the core.

In the picture the black lines represent the default links: the fixed connections among the network elements. The other

lines instead are the dynamically allocated cut-through paths. Those follow the traffic characteristics and can either be terminated in one of the backbone routers or else remain optical up to the border routers (depending on the wavelength capabilities of the latter).

One wavelength is persistent and statically allocated, allowing the IP layer to exchange routing information following existing algorithms (OSPF, BGP). The same channel is also used to send data traffic and the signalling messages for dynamically allocating new cut-through paths.

A distinctive advantage of the architecture we introduce is that it is completely IP-based: the default connections may actually represent an existing IP network, while dynamically allocated paths are based on recognition of IP flows.

The optical physical layer can be based either on SDH/Sonet or Gigabit (and 10G) Ethernet. The use of both the technologies together limits the functionalities of the router, as optical paths created with one type of interface will not be extendable to the other. On the other hand using the same technology both at the edge and in the backbone of the network will simplify the overall network design and reduce the cost of implementation. Ethernet is ideal because of its low level of complexity which helps simplifying the implementation and extending the optical cut-through paths up to the end-user routers. Moreover it can scale from 10 Mbit to 10Gbit per second and both the switches and interfaces prove to be very cost effective and are widely available [18] (prices range from less than $50 for a copper-based interface to less than $500 for a fiber-based interface). On the other hand however optical Ethernet has not yet got all the transport capabilities of the SDH/Sonet transport plane (e.g. controlled signal regeneration, per-channel data rate up to 40G).

A first practical implementation would probably see both the interfaces working on the routers, while as its technology improves Ethernet may remain the only interface.

## 3    System implementation:

Figure 3 represents our implementation of the photonic router. The system is made up of a modified IP router with multiple Gigabit Ethernet (and 10G) interfaces, directly connected to a photonic switch. The switch, which is directly controlled by the IP router, connects the whole system to the external world by multiple WDM links which are demultiplexed before entering the switch.

After the initialization phase, where the router runs a self-configuration algorithm to discover its ports' connectivity, the neighbour discovery phase is initiated over the standard links. Once this configuration phase is over, the router is ready to process traffic and dynamically allocate new wavelengths.

All the packets coming from the default interfaces are analyzed by the flow recognition algorithm. Once a flow with predetermined characteristics is encountered, the node will signal its immediate upstream and downstream neighbours to

find a common wavelength where the cut-through path could be established. Once a common wavelength is found, the router will instruct the photonic switch to create a direct optical link between the upstream and the downstream node. This process can be speeded up if each router is constantly updated on the status of the resources of its direct neighbours, eliminating the latency time due to multiple acknowledgments.
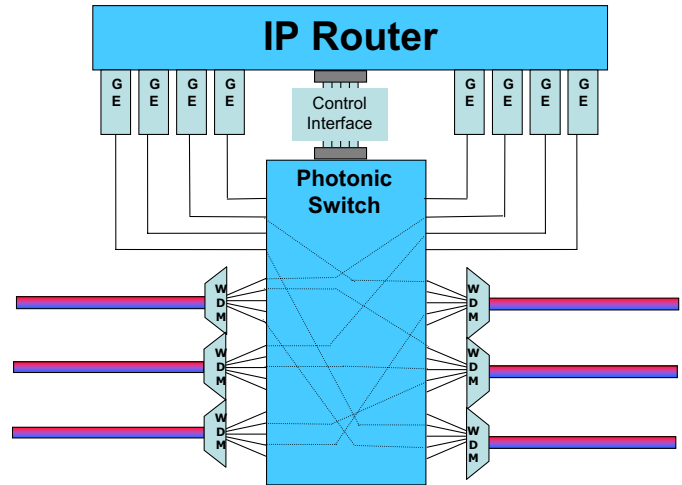


**Figure 3 Diagram with the essentials of an IP-switched optical router.**

The source of the new trail (intended as the upstream neighbour of the switch that has detected the flow, not the IP source of the flow) has the task of injecting the packets belonging to the flow into the dedicated path. It is also in charge of controlling the flow, by sending periodical refresh messages along the default links to keep reserving the path, or sending a reset message when the flow is no longer active.

One of the main functionalities of the system is the possibility to extend the flow to more than 3 nodes. Figure 4 shows an example of path extension: the light trail, initially established between 3 nodes, is extended to the last node which becomes the new destination of the path.

This concept brings a lot of new issues and challenges to the project. First of all there is a trade-off between how far the path can be extended and how many flows can be aggregated into the same path. In fact as the path length increases (measured by the number of hops) it will be more and more difficult to find suitable flows with a long common path.

Other difficulties will arise from the fact that the system is completely distributed. For example the same flow could be assigned different wavelengths in different parts of the network. If the lifetime of the allocated path is long enough, the extension mechanism may create a situation where the same node becomes both the destination of one part of the flow, working on a certain wavelength, and the source of the other part of the flow, which may run on a different wavelength. At this point the joining of the two paths could

easily be achieved if that switch has wavelength conversion capabilities. A more complicated solution would require signalling a part of the network for a change of wavelength: the feasibility of such approach will be distributedly determined by the nodes.
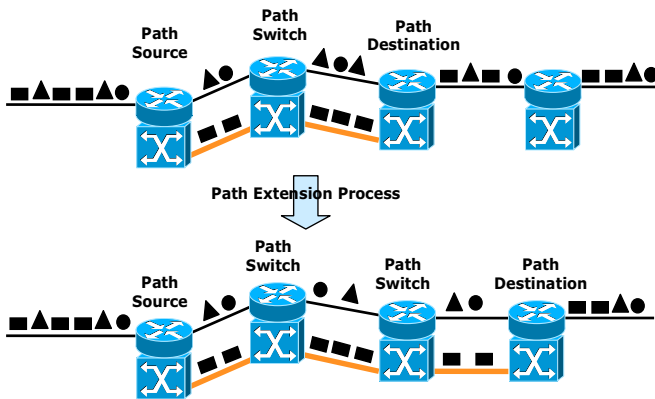


**Figure 4 Diagram illustrating the path extension process.**

Not less challenging will be the study of new algorithms to select those flows that would most conveniently exploit the available optical bandwidth. In fact among the high number of encountered flows the router has to choose those which are more likely to bring the highest percentage of traffic, more prone to be extended and with higher life time expectations (compared to the switching time of the photonic switch).

## 3.1 Optical test bed:

The architecture of our testbed is displayed in Figure 5.
The network we have implemented is made up of 6 nodes, connecting the CTVR laboratory in Trinity College with the Optical Communication Laboratory in Dublin City University, over a 16 Km link.
The three routers are connected to each other and one end user machine is connected to each router.
The connection between the two universities runs over a three-wavelength dedicated bi-directional optical link, supplied by HEAnet (Ireland's National Education and Research Network). One of the lambdas is used as a default connection, while the remaining two are used to dynamically accommodate the optical cut-through paths.
All the optical links are based on Gigabit Ethernet and the photonic switches are MEMs 16x16-port Glimmerglass systems.
The routers are blade servers and the protocol stack we have developed runs on top of Windows 2003 OS. While this hardware would not cope with any significant packet load, it has the advantage of simplifying the protocol development. In this phase we are focusing on demonstrating and testing the concept: the maximization of the throughput will be addressed in a further stage with an FPGA implementation. The optical testbed will allow us to study and develop solutions for issues related both to the signalling mechanisms to create and extend

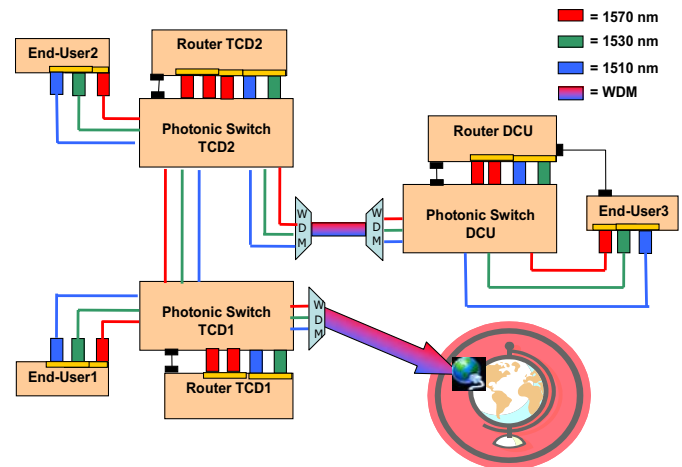the flows and to the flow analysis algorithm to trigger the path creation.



**Figure 5 Testbed developed at CTVR.**

Currently we have a first version of the signalling protocol implemented and a simple flow recognition algorithm, based on the data rate estimation of the encountered flows. This permits us already to automatically allocate optical cut-through paths depending on the data rate of the incoming flows.

This implementation however does not use a real IP layer for packets forwarding. In fact we are studying the behaviour of more dynamic protocols like those typically used in ad-hoc networks (e.g. DSR, OLSR). Nevertheless the only information we use to create the dynamic paths is the source and destination addresses of the packets. This makes our system fully compatible with the IP layer and with any network layer which carries this type of information in its header.

Finally, in the figure the Photonic Switch labelled TCD1 has some open connections that may allow us to connect our network to other existing research networks and testbeds available internationally. In this way we may be able to interface to testbeds taking different approaches.

## 4 Conclusions:

In this paper we have presented an optical testbed consistent with the idea of disaggregated network architecture. In contrast to the GMPLS standard, where the optical paths are end-to-end and requested by a user, our optical light paths are locally created and triggered by the traffic encountered at the optical routers.
This novel architecture reveals many open challenges, from flow recognition to path extension and aggregation. The testbed will be used to examine solutions to those issues and produce results that will allow comparison with other existing network architectures.

# References:

[1] D. J. Arent, A. Martin. Third-generation DWDM networks near reality, Lightwave March 2001.

[2] G. Bernstein, B. Rajagopalan, D. Saha. Optical Network Control. Addison-Wesley, June 2003.

[3] J.D. Jones, L. Ong, M.A. Lazer. Creating an Intelligent Optical Network Worldwide Interoperability Demonstration. Communications Magazine, IEEE, Volume 42, Issue 11, Nov. 2004.

[4] D. Saha, B. Rajagopalan, G. Bernstein. The Optical Network Control Plane: State of the Standards and Deployment. Communications Magazine, IEEE , Volume: 41 , Issue: 8 , Aug. 2003.

[5] L. Berger *et al.* Generalized MPLS – RSVP-TE Extensions. IETF RFC 3473, Jan. 2003.

[6] P. Ashwood-Smith *et al*. Generalized MPLS - CR-LDP Extensions. IETF RFC 3472, Jan. 2003.

[7] Reconnected to grow: Global Telecommunications Industry Index 2005. Published by Deloitte Touche Tohmatsu, 2005.

[8] P. Newman, G. Minshall, T.L. Lyon. IP Switching-ATM Under IP. Networking, IEEE/ACM Transactions on , Volume: 6 , Issue: 2 , April 1998.

[9] S. Lin, N. McKeown. A Simulation Study of IP Switching. Proceedings of the ACM SIGCOMM '97 conference. Cannes, France.

[10] M. Kuznetsov, N. M. Froberg, S.R. Henion, H.G. Rao, J. Korn, K.A. Rauschenbach, E.H. Modiano, V.W.S. Chan. A Next-Generation Optical Regional Access Network. Communications Magazine, IEEE Volume 38, Issue 1, Jan. 2000.

[11] M. Veeraraghavan, R. Karri, T. Moors, M. Karol, R. Grobler. Architectures and Protocols that Enable New Applications on Optical Networks. Communications Magazine, IEEE, Volume 39, Issue 3, March 2001.

[12] K. Sato, N. Yamanaka, Y. Takigawa, M. Koga, S. Okamoto, K. Shiomoto, E. Oki, W. Imajuku. GMPLS-Based Photonic Multilayer Router (Hikari Router) Architecture: An Overview of Traffic Engineering and Signaling Technology. Communications Magazine, IEEE, Volume: 40 , Issue: 3 , March 2002.

[13] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku, Y. Takigawa. Dynamic Multilayer Routing Schemes in GMPLS-Based IP + Optical Networks. Communications Magazine, IEEE , Volume: 43 , Issue: 1 , Jan. 2005.

[14] W. Willinger, V. Paxson, M. Taqqu. A Practical Guide to Heavy Tails: Statistical Techniques and Applications, chapter Self-similarity and Heavy Tails: Structural Modeling of Network Traffic, pages 27–54. R. Adler, R. Feldman, and M. S. Taqqu, editors. Birkh¨auser Verlag, Boston, 1998.

[15] F. Hernandez-Campos, J.S. Marron, G. Samorodnitsky, F.D. Smith. Variable Heavy Tailed Durations in Internet Traffic. Part I: Understanding Heavy Tails. Modeling, Analysis and Simulation of Computer and Telecommunications Systems. Proceedings. 10th IEEE International Symposium on 11-16 Oct. 2002.

[16] C. Qiao, M. Yoo. Optical burst switching (OBS): A new paradigm for an optical Internet. Journal of high speed networks, 8(1), 1999.

[17] Y. Xiong, M. Vandenhoute, H.C. Cankaya. Control architecture in optical burst-switching WDM networks. IEEE Journal on Selected Areas in Communications, Oct. 2000.

[18] The Metro Ethernet Network: Comparison to Legacy SONET/SDH MANs for Metro Data Service Providers. Metro Ethernet Forum Whitepaper, July 2003.